

**The Relationship between Gradability and Prenominal Adjective Ordering
Preferences**

Denny O'Shea

Linguistics Department

San Jose State University

LING202: Syntactic Theory

Yining Nie

April 19, 2023

The Relationship between Gradability and Prenominal Adjective Ordering Preferences

Generative grammar aims to identify cross-linguistic patterns in syntax, but this endeavor is far from straightforward. Syntax is intricately intertwined with semantics and morphology, and interdisciplinary discussions are often necessary to fully understand the nature of linguistic universals. One area that has garnered a great deal of attention and may contain an interdisciplinary answer is the preference for certain orders of adjectives in noun phrases. Linguists such as Cinque have long recognized a degree of regularity across languages in how adjectives are ordered. In Cinque (1994), he notes that there appears to be no unmarked variation in the ordering of adjectives prenominally. He hypothesizes that adjectives are placed in a strict series of maximal projections. He does not, however, explain what motivates that specific ordering in the first place. Given that this is a strong claim, it is important to understand the potential psychological motivations for this cross-linguistic phenomenon.

Several hypotheses have been proposed, with most agreeing on a correlation between the semantics of both the noun and the modifiers (Scontras, 2023; Scontras et al., 2017). A particularly noteworthy study in this area was conducted by Scontras et al. (2017). Through experimentation and corpus validation, Scontras et al. (2017) argued that the hierarchy of maximal projections proposed by Cinque (1994) is motivated by the subjectivity of adjectives. In their study, they found a spectrum of subjectivity with the least subjective adjectives closer to the noun than the most subjective (Scontras et al., 2017). However, what is not clear is why subjectivity would correlate so distinctly between semantic classes. As is well documented, adjectives are inherently vague and thus subjective (Kennedy, 2007; Kennedy & McNally, 2005). What is not inherent in all adjectives, however, is gradability.

Adjectives are said to be gradable if they can be modified to some degree, and if that modification is linked to some standard of comparison (Kennedy & McNally, 2005).

This standard is relative to the external factors governed in the realm of discourse. For instance, a phrase like (1a) below can mean nothing on its own. It is only when the relevant context in (1b) is provided that readers can make any sort of judgment about its truth parameters (Kennedy, 2007).

- (1) a. The coffee in Rome is expensive.
- b. The coffee in Rome is more expensive than in Chicago.

As Kennedy (2007, p. 3) notes, this is due to the relationship between the gradability of the adjective “and a context dependent standard of comparison.” Thus, any measurement of these adjectives’ subjectivity must consider the effect context will have. Scontras (2023, p. 362) briefly mentions that Cinque acknowledged this context dependence when constructing his hypothesis of maximal projections. However, the extent to which the standard of comparison affects preferred adjective ordering has yet to be explored.

The key factor in distinguishing between gradable and non-gradable adjectives is the standard of comparison (Kennedy, 2007; Solt, 2016). Solt (2016) also argues that subjectivity is speaker-dependent and, thus, scalable based on the speaker’s perception of a situation. For instance, an adjective such as *pretty* can compare two people with degrees of modification, but the order in which someone might place those people depends on the speaker’s perception of the adjective. That is to say, Speaker A may think that Person A is prettier than Person B, but Speaker B may disagree given their own standard of comparison.

The two main categories of gradability—**relative** and **absolute**—can be further differentiated based on their use in discourse. Relative gradable adjectives are context-dependent adjectives (Kennedy & McNally, 2005) that can be identified by their use in *for* phrases and inability to occur with “end-point oriented degree modifiers such as *slightly* and *completely*,” according to Solt (2016, p. 676). On the other hand, absolute adjectives are not context-dependent, scalable with minimum and maximum standards, and can be modified using degree modifiers. Non-gradable adjectives cannot be compared

nor used with degree modifiers. Their relationship to context, however, is not clear. Solt (2016) also identifies a third category of gradability which is *evaluative*. These adjectives are highly subjective, depend on a speaker’s “predicates of personal taste,” and are easiest to identify (Solt, 2016, p. 680).

Solt (2016) shows that subjectivity can fluctuate based on the type of adjective present. Based on this research, it is interesting that Scontras et al. (2017) would suggest that subjectivity is so clear cut between semantic classes. Is it possible that the measurement proposed by Scontras et al. (2017) is simply an abstraction of the degrees of gradability? Furthermore, could these degrees of gradability be alternating in the same patterns they do in Solt (2016), thus motivating the strict hierarchy proposed by Cinque (1994)?

Faultless Disagreement of Gradable Adjectives

The results of Scontras et al. (2017) rely on the assumption that subjectivity is kept constant by speakers. However, research on this subject does not suggest this is the case (Kennedy, 2007; Kennedy & McNally, 2005; Solt, 2016). This is one of the major oversights of the Scontras et al. (2017) study, as their faultless disagreement test consists of various adjectives tested against only one constant noun. As our discussion above suggests, subjectivity is based on some standard of comparison Kennedy (2007). Thus, one cannot make a reasonable assumption about an adjective’s subjectivity without giving context. Words that seem subjective based on abstract ideas may turn out to be more concrete with context.

Participants

This survey recruited 25 participants through word-of-mouth, Facebook advertisements, and survey links posted to Reddit forums designated by SurveySwap.io for students seeking survey respondents. Participants were assumed to be native English speakers, as all three outreach programs targeted this demographic. However, word-of-mouth recruits may have included some highly competent second-language

speakers. The survey was anonymous, with no demographic information requested to screen for native proficiency. Participants were not compensated for their time outside of tokens valid for SurveySwap.io. This program helps students find survey participants by purchasing outreach tokens through cash or completing other surveys on their site.

Materials and Design

Stimuli comprised the 10 nouns and 67 adjectives presented in Scontras et al. (2017, p. 61). One or two adjectives were excluded from the original set due to their incongruity with the task. All words in the set were semantically categorized, and adjectives were given the additional category of gradability. Furthermore, to ensure that combinations were not completely random, a third animate category was designated so that adjectives deemed only appropriate with animate objects would not be paired with inanimate objects.

Each adjective was categorized for gradability using the tests in Solt (2016). Here, relative gradable adjectives are those with “context- or comparison class-dependent standards”, denoted by their ability to occur in *for*-phrases and without “endpoint-oriented degree modifiers” (Solt, 2016, p. 1). Absolute gradable adjectives are those which “have scalar maxima or minima as standards” and can be used with endpoint-oriented degree modifiers. Non-gradable adjectives are limited in their use in *for*-phrases and with modifiers, and evaluative adjectives are limited by neither. This last category is deemed the most subjective as they are value judgements and dependent on a speaker’s “predicated of personal taste” (Solt, 2016, p. 2). The results of these tests and the final classification of each adjective and noun can be found in Appendix A.

Once classification was completed, a Python script was designed that utilized a random number generator and the Pandas sample method (McKinney, 2010) to extract noun/adjective combinations. First, a sample of adjectives was collected to ensure a distribution of both semantic and gradability categories similar to the original set of stimuli. Each adjective was then checked for animacy before a noun was selected. If the adjective was deemed only appropriate with animate nouns, the pool consisted of only five

animate nouns. If not, the pool was all 10 nouns. Each noun/adjective pair was then placed in a template like this:

(2) Mary thinks this mouse is sadder than that mouse.

John doesn't think so.

Can they both be right or must one be wrong?

All the noun/adjective combinations were saved in a Google Sheet and transferred to Google Forms with the Form Builder for Sheets extension. The final output of the survey consisted of 46 unique noun/adjective combinations presented in the (2) format with a binary response for each. A “Yes” response indicates that faultless disagreement is permissible in this context; a “No” indicates this is a factual disagreement. Each response was tallied and organized in a DataFrame for further analysis.

Hypothesis

As stated before, subjectivity is inextricably tied to the standard of comparison provided by context (Kennedy, 2007). Given that the faultless disagreement test in this task does provide a context not present in the Scontras et al. (2017), I hypothesize that the rate of faultless disagreement will produce a slightly different adjective order than that presented in the original paper. Furthermore, Solt (2016) shows that while evaluative adjectives are highly subjective with context, relative and absolute gradable adjectives do not follow a strict subjectivity pattern but alternate on a word-by-word basis. I should expect that the results of this faultless disagreement task will have the same outcome with semantic categories. While Solt (2016) does not test non-gradable adjectives, I should expect them to perform parallel to evaluative adjectives and be highly non-subjective even with the given context.

Results

The results of this experiment are available in Figure 1. As stated, each response was tallied, with “Yes” responses indicating a preference for faultless disagreement, and “No” responses indicating a factual disagreement preference. In order to generate the bar

graph depicted in Figure 1, a Python script was utilized to compute the mean percentage of “Yes” responses for a particular semantic category. Seaborn (Waskom, 2021) was employed to produce the graph.

Discussion

As expected, evaluative adjectives were the most subjective of the gradability types, being categorized by subjects as a case of faultless disagreement 89.71% of the time on average. Surprisingly, however, non-gradables in this task had the second highest percentage on average at 54.10%. Relative and absolute gradables

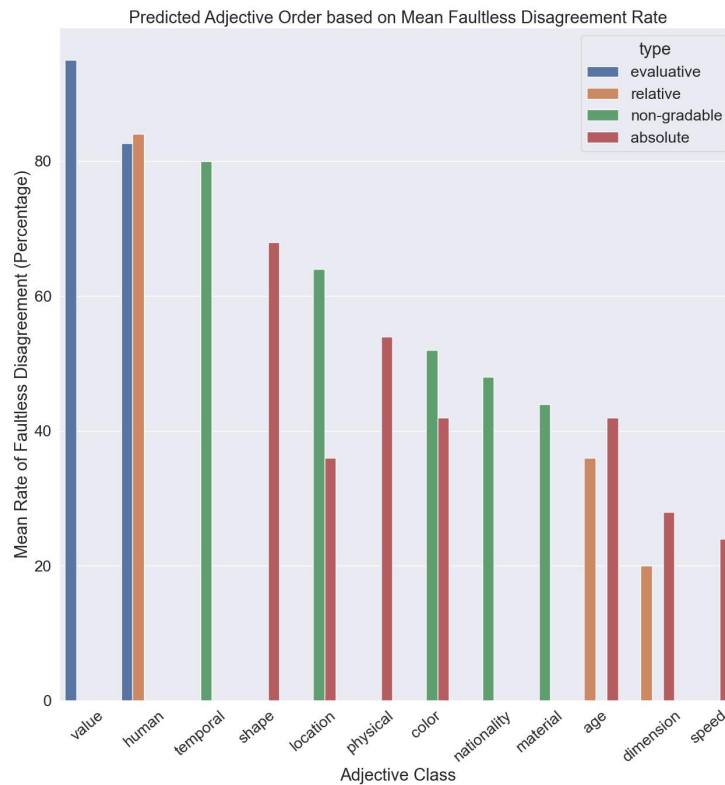


Figure 1

Average rating of faultless disagreement per semantic category.

showed the lowest rates of faultless disagreement on average, with 46.67% and 42.57% average ratings respectively. These results, while following a pattern of alternation similar to that in Solt (2016), do not suggest that gradability is inherently linked to subjectivity. It also does not suggest subjectivity and context are unrelated. For this, the semantic categories are useful.

The data are organized by averaging the percentage of "Yes" responses for each semantic category. The order in which the categories are displayed should indicate whether subjectivity is inherent or context-dependent. If semantic categories are inherently

subjective whether or not context is available, then the order of these categories should be the same as in Scontras et al. (2017). However, this is not the case. Most notably, the semantic category deemed most subjective in Scontras et al. (2017), dimension, is actually one of the least subjective with the addition of a standard of comparison. Perhaps more notable for the scope of this research is that the two categories in the middle of the Scontras et al. (2017) subjectivity spectrum, age and physical, are now reversed with a 14% difference in faultless disagreement ratings on average.

Word Order Flexibility Task

Scontras et al. (2017) argues that prenominal adjective ordering is motivated by the semantic category of an adjective and the inherent subjectivity of that category. However, as the previous experiment shows, rates of subjectivity between these categories are not inherent and change with the addition of a standard of comparison. This suggests that the Scontras et al. (2017) assumption that motivation for prenominal adjective order is based on subjectivity does not seem to hold when controlling for this factor. The question, then, of why prenominal adjectives prefer a specific word order remains unclear. This second experiment is designed with this question in mind.

Participants

This second survey recruited 27 participants through Facebook advertisements, and survey links posted to Reddit forums designated by SurveySwap.io for students seeking survey respondents. Just as in the previous experiment, participants were assumed to be native English speakers. The survey was anonymous, with no demographic information requested to screen for native proficiency. Again, participants were not compensated for their time outside of tokens valid for SurveySwap.io.

Materials and Design

Scontras et al. (2017) suggest a connection between semantic categories and a spectrum of subjectivity, promoting the idea that Cinque’s strict hierarchy of prenominal adjective ordering is motivated by this spectrum. However, the results of their research

show an insignificant difference in the mean distance from the noun between the semantic classes *age* and *physical*. This is particularly true of their faultless disagreement and subjectivity experiments. Furthermore, the results of my faultless disagreement experiment show a reversal of these two classes. This supports the proposal that perhaps there is some level of flexibility permissible between these two semantic categories not allowed by Cinque’s strict hierarchy.

To explore this, I employed the same methods as Scontras et al. (2017) in their exploration of mean word order. For this task, participants were asked to identify which of two descriptions sounded more natural by rating their preference on a 5-point scale. Ratings of 3 indicated no preference for the Cinque and Scontras suggested adjective order nor the alternative order; ratings greater than 3 indicated a preference for the Scontras order; ratings less than 3 indicated a preference for the alternate order. Stimuli were again pulled randomly from a list of nouns and adjectives, which were pulled from the materials section of Scontras et al. (2017) and can be found in Table 1.

Table 1

Adjectives, nouns, and their semantic classes

Word	Category	Class	Word	Category	Class
apple	food	noun	new	age	adjective
banana	food	noun	old	age	adjective
carrot	food	noun	rotten	age	adjective
cheese	food	noun	fresh	age	adjective
tomato	food	noun	smooth	physical	adjective
chair	furniture	noun	hard	physical	adjective
couch	furniture	noun	frozen	physical	adjective
fan	furniture	noun	soft	physical	adjective
desk	furniture	noun	sweet	physical	adjective

Once samples were collected, they were screened for duplicates, nonsense

combinations, and high collocation rates. The final output of the survey appeared something like in Figure 2 below.

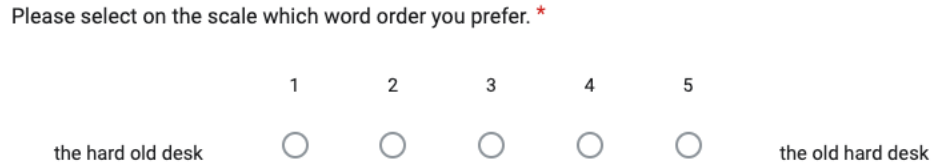


Figure 2

Example trial from Word Order Experiment: Participants indicated which of the two adjective combinations sounds more natural on a 5-point linear scale.

Hypothesis

Scontras et al. (2017) concludes that subjectivity alone is what motivates the Cinque hierarchy. If this is the case, then those semantic categories in between the two extremes of subjectivity should show higher rates of word order flexibility. However, if it is the case that gradability plays a greater role than previously assumed, then it would not follow that flexibility is possible. While the previous experiment does not suggest that the hierarchy is motivated by gradability alone, it does show an alternating pattern where two categories of gradability are rarely side-by-side. While the two semantic categories generally do belong to the same gradability class, there may be other categories in between not present in the Scontras et al. (2017) data blocking movement. Thus, their placement in the hierarchy would be based on the finer detail of gradability and not on their respective subjectivity.

Results

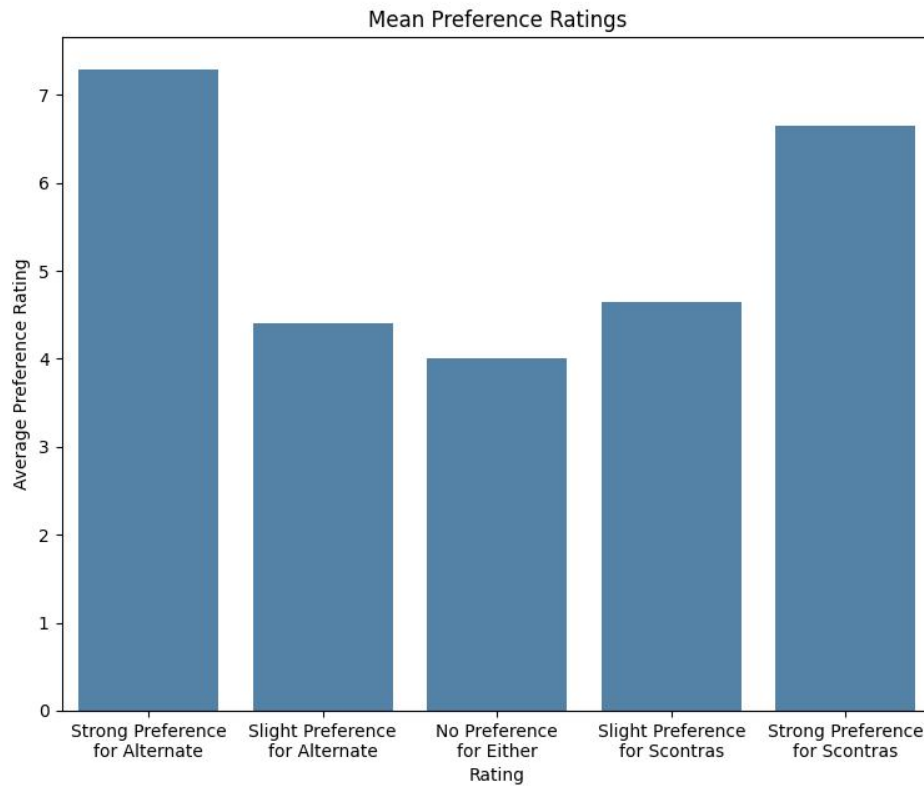


Figure 3

Results of Experiment 2: Mean preference ratings reported by subjects on the naturalness of the Cinque analysis versus the alternative order.

Participants rated their preference for one order over the other on a 5-point linear scale, with ratings above 3 indicating a preference for the adjective order suggested by Scontras and Cinque, ratings below 3 indicating a preference for the alternative order, and a rating of 3 indicating no preference. The mean preference ratings are presented in Figure 3.

To analyze the statistical significance of the differences in preference between the categories, a Kruskal-Wallis test was performed. The Kruskal statistic was calculated as 4.428, with a corresponding p-value of 0.109. The non-significant p-value suggests that there is no significant difference in word order preference between the semantic categories.

Discussion

These findings, along with the results of the first experiment, show that gradability does not seem to play a role in the hierarchy of prenominal adjective ordering. Instead, it further supports the Scontras et al. (2017) assumption that subjectivity plays a pivotal role as these are the categories within the middle of their spectrum. Thus, flexibility in which one comes first is generally expected according to these results.

General Discussion

There has been significant attention to prenominal adjective order preferences in the last several decades. Cinque has long argued for a strict hierarchy of maximal projects to explain the cross-linguistic preference for one unmarked order of prenominal adjectives. However, exactly what motivates this hierarchy is not yet clear. Scontras et al. (2017) posits that it is due to the perceived inherent subjectivity of particular semantic classes that is the motivating factor.

The research in this paper explores this notion more thoroughly, paying particular attention to the assumption that certain semantic classes will be perceived as subjective at a constant rate. Previous research does not suggest this is the case, especially when given a standard of comparison as context. The data in this paper further supports this, showing that rates of subjectivity for adjective classes can change dramatically based on the context present. Thus, if one were to assume that all prenominal adjectives follow a strict hierarchy, it is not clear why subjectivity would be the sole motivating factor.

However, this argument also assumes the premise that the exact order of semantic classes is well documented and understood, which is not the case (Scontras, 2023). In fact, the results in Scontras et al. (2017) do not support a strict hierarchy. Experiment two of this paper shows permissible flexibility between the classes in the middle of their subjectivity spectrum, *age* and *physical*. This follows, given that there is no significance between these two classes in the Scontras et al. (2017) experiments for subjectivity and faultless disagreement. Of course, this is an ongoing investigation. Further research on a

larger scale may be needed to make definitive conclusions about the complex interplay between subjectivity, gradability, and prenominal adjective ordering.

References

- Cinque, G. (1994). On the evidence for partial N-movement in the Romance DP. In R. S. Kayne, G. Cinque, J. Koster, J. Y. Pollock, L. Rizzi, & R. Zanuttini (Eds.), *Paths toward universal grammar: Studies in honor of Richard S. Kayne* (pp. 85–110). Washington, DC: Georgetown University Press.
- Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguist Philos*, (30), 1–45.
<https://doi.org/https://doi.org/10.1007/s10988-006-9008-0>
- Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, 81(2), 345–379.
- McKinney, W. (2010). Data structures for statistical computing in Python. *Proceedings of the 9th Python in Science Conference*, 56–61.
<https://conference.scipy.org/proceedings/scipy2010/mckinney.html>
- Scontras, G. (2023). Adjective ordering across languages. *Annual Review of Linguistics*, 357–373. <https://doi.org/https://doi.org/10.1146/annurev-linguistics-030521-041835>
- Scontras, G., Degan, J., & Goodman, N. D. (2017). Subjectivity predicts adjective ordering preferences. *Open Mind: Discoveries in Cognitive Science*, 1(1), 53–65.
https://doi.org/https://doi.org/10.1162/opmi_a_00005
- Solt, S. (2016). Ordering subjectivity and the absolute/relative distinction. *Proceedings of Sinn und Bedeutung*, 20, 676–693.
<https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/289>
- Waskom, M. (2021). Seaborn: Statistical data visualization. <https://seaborn.pydata.org>

Appendix
Stimuli Used for Faultless Disagreement Task

Word	Category	Animate	Type
junior	age	TRUE	absolute
senior	age	TRUE	absolute
black	color	FALSE	absolute
white	color	FALSE	absolute
big	dimension	FALSE	absolute
large	dimension	FALSE	absolute
long	dimension	FALSE	absolute
narrow	dimension	FALSE	absolute
open	dimension	FALSE	absolute
thick	dimension	FALSE	absolute
thin	dimension	FALSE	absolute
close	location	FALSE	absolute
creamy	physical	FALSE	absolute
curly	physical	FALSE	absolute
frozen	physical	FALSE	absolute
smooth	physical	FALSE	absolute
spicy	physical	FALSE	absolute
sweet	physical	FALSE	absolute
circular	shape	FALSE	absolute
fast	speed	FALSE	absolute
slow	speed	FALSE	absolute

Continued on next page

Word	Category	Animate	Type
civilized	human	TRUE	evaluative
creative	human	FALSE	evaluative
entrepreneurial	human	TRUE	evaluative
playful	human	TRUE	evaluative
professional	human	TRUE	evaluative
sad	human	FALSE	evaluative
selfish	human	TRUE	evaluative
speedy	speed	FALSE	evaluative
best	value	FALSE	evaluative
exciting	value	FALSE	evaluative
lavish	value	FALSE	evaluative
plain	value	FALSE	evaluative
pleasant	value	FALSE	evaluative
prestigious	value	FALSE	evaluative
strange	value	FALSE	evaluative
blonde	color	TRUE	non-gradable
blue	color	FALSE	non-gradable
green	color	FALSE	non-gradable
purple	color	FALSE	non-gradable
red	color	FALSE	non-gradable
yellow	color	FALSE	non-gradable
internal	location	FALSE	non-gradable
overhead	location	FALSE	non-gradable
corduroy	material	FALSE	non-gradable

Continued on next page

Word	Category	Animate	Type
crocheted	material	FALSE	non-gradable
gold	material	FALSE	non-gradable
wooden	material	FALSE	non-gradable
international	nationality	TRUE	non-gradable
national	nationality	TRUE	non-gradable
Brazilian	nationality	FALSE	non-gradable
English	nationality	FALSE	non-gradable
European	nationality	FALSE	non-gradable
Hispanic	nationality	FALSE	non-gradable
Japanese	nationality	FALSE	non-gradable
Vietnamese	nationality	FALSE	non-gradable
solid	physical	FALSE	non-gradable
square	shape	FALSE	non-gradable
current	temporal	FALSE	non-gradable
daily	temporal	FALSE	non-gradable
everyday	temporal	FALSE	non-gradable
historical	temporal	FALSE	non-gradable
new	age	FALSE	relative
old	age	FALSE	relative
old-time	age	FALSE	relative
young	age	FALSE	relative
strict	human	TRUE	relative